

**REMARKS**

In the Office Action, claims 1-43 were finally rejected under 35 U.S.C. § 102(e) as allegedly being anticipated by Application Publication No. 2001/0000962 of Rajan (hereinafter "Rajan").

Applicants respectfully traverse the rejections of record, and further submit that claims 1-43 are in condition for allowance.

Rejections under 35 U.S.C. § 102(e) in view of Rajan

Claims 1-43 were rejected under 35 U.S.C. § 102(e) as allegedly anticipated by Rajan.

Independent claim 1 is directed to a system for generating a description record from multimedia information, comprising, *inter alia*:

a computer processor, coupled to said at least one multimedia information input interface, receiving said multimedia information therefrom, processing said multimedia information by performing *object extraction processing* to generate multimedia object descriptions from said multimedia information, and processing said generated multimedia object descriptions by *object hierarchy processing* to generate multimedia object hierarchy descriptions indicative of an organization of said object descriptions, wherein *at least one description record including said multimedia object descriptions and said multimedia object hierarchy descriptions is generated* for content embedded within said multimedia information

Independent claim 17 includes similar limitations in the context of a method claim.

Rajan is directed to a method and apparatus for composing and presenting multimedia programs using the MPEG-4 standard at a multimedia terminal, including an

architecture wherein the composition of a multimedia scene and its presentation are processed by two different entities – a “composition engine” and a “presentation engine.” *See* Rajan, ¶ 0002. In this sense, the description of Rajan is limited to the MPEG-4 standard, “which allows a user to interact with video and audio objects within a scene,” and which allows a user to modify scenes by deleting, adding, or repositioning objects, or changing the characteristics of objects, such as size, color, and shape, for example. *See* Rajan, ¶ 0004. Rajan is thus directed to a different problem, i.e., composing and presenting multimedia video, from that of the present invention, i.e., techniques for describing multimedia information content to enable intelligent searching of multimedia content via, e.g., the Internet. *See* Specification, p.1 , lines 1-4, p. 9, lines 23-29.

Another essential difference is the definition of “object.” Rajan merely uses “object” to refer to portions of media content, for example, an image region. In the present invention, “object” refers not only to a portion of some media content but also, importantly, objects in the real world that are depicted in the multimedia content. In addition, an “object” in Rajan can only have content of one media type. However, the present invention is generic to consider also portions of multimedia documents including different media.

The Examiner, on pp. 4-6 of the Final Office Action, maintains that ¶¶ 0042 – 0046 of Rajan disclose all elements of claim 1. Applicant respectfully disagrees. Not only do these paragraphs *not* disclose all elements of the claim, they are not even directed to solving the same problem as the claimed invention, and, moreover, do not even relate to the same subject matter.

Rajan does not disclose or suggest “processing said multimedia information by

*performing object extraction processing* to generate multimedia object descriptions.” The Examiner alleges that ¶ 0042 of Rajan discloses the claimed object extraction. However, ¶ 0042 of Rajan states:

According to the MPEG-4 Systems standard, the scene description information is coded into a binary format known as BIFS (Binary Format for Scene). This BIFS data is packetized and multiplexed at a transmission site, such as a cable and or satellite television headend, or a server in a computer network, before being sent over a communication channel to a terminal 100. The data may be sent to a single terminal or to a terminal population. Moreover, the data may be sent via an open-access network or via a subscriber network.

This portion of Rajan is directed to BIFS (Binary Format for Scene) coding, as set forth in the MPEG-4 standard. The BIFS data is packetized and transmitted over a communications link. Applicant does not see *any relation* between this paragraph and object extraction for generating multimedia descriptions, as claimed in the present invention.

Likewise, ¶ 0043 is cited as allegedly disclosing the claimed feature of “*processing said generated multimedia object descriptions by object hierarchy processing* to generate multimedia object hierarchy.” However, while ¶ 0043 generally discloses that an “MPEG-4 scene follows a hierarchical structure”, it nowhere indicates that multimedia objection descriptions, which are generated by, e.g., performing object extraction, are processed by object hierarchy processing.

Furthermore, the claimed hierarchy and the hierarchy of Rajan are completely different. The cited paragraph of Rajan, ¶ 0043, states that “an MPEG-4 scene follows a

hierarchical structure.” Notably, ¶ 0043 contains the *only* reference to hierarchy in Rajan. This hierarchy is not described with sufficient detail, and certainly is not enabling to one of ordinary skill in the art. For at least this reason, the claimed “processing said generated multimedia object descriptions by object hierarchy processing to generate multimedia object hierarchy” is not disclosed for purposes of anticipation.

Moreover, the claimed hierarchy is apparently entirely different from that of the present invention. The cited paragraph of Rajan, ¶ 0043, states that “an MPEG-4 scene follows a hierarchical structure.” This would suggest that the “tree structure” discussed in Rajan is for purposes of flow of a video scene in space and time. This is further apparent when considered in the context of the problem which Rajan is directed to, i.e., *composing* and *presenting* multimedia video *in space and time*. However, the claimed hierarchy, as further described, e.g., at p. 13 of the present application, relates to an object hierarchy for description of particular video objects with varying levels of specificity – for purposes of content description, and *not* hierarchy of a scene for composing or presenting the scene. In addition, the hierarchy of the present invention is not only based on temporal and spatial constraints, but may also include visual, motion and semantic information. This is further apparent by reference to a comparison of MPEG-4 and MPEG-7 standards papers. For the Examiner’s reference, Application directs the Examiner to:

<http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm>

<http://www.chiariglione.org/mpeg/standards/mpeg-4/mpeg-4.htm>

which provide further detail regarding MPEG-4, MPEG-7, and the differences in the hierarchical structures of each.

Accordingly, because Rajan fails to disclose or suggest at least these claimed features, and is in fact related to an entirely different object and invention, this reference fails to anticipate independent claims 1 and 17. Applicants respectfully submit that these claims are in condition for allowance. Additionally, because all depending claims 2-16 and 18-32 contain the foregoing limitations through dependency from claims 1 and 17, Applicants respectfully submit that these claims are also in condition for allowance.

Claims 3, 7, 19 and 23 are patentable for the additional reason that they include limitations of “image segmentation” and “feature extraction.” The Examiner asserts that ¶ 0045 of Rajan, copied below, describes the claimed features:

The scene description information can also indicate attribute value selection. Individual media objects and scene description nodes expose a set of parameters to a composition layer through which part of their behavior can be controlled. Examples include the pitch of a sound, the color for a synthetic object, activation or deactivation of enhancement information for scaleable coding, and so forth.

Rajan, ¶ 0045.

Applicant again cannot find any reference in the above-cited paragraph to the claimed “image segmentation” and “feature extraction.” Indeed, there is none. Accordingly, for at least this additional reason, claims 3, 7, 19 and 23 are further patentable over Rajan.

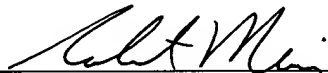
Independent claim 33 is directed to a computer readable medium with at least one multimedia description record describing multimedia content for corresponding multimedia information, the description record comprising, *inter alia*, object descriptions, generated by performing object extraction processing. The Examiner has rejected claim 33 for the same reasons as claims 1 and 17. However, as discussed at length above, Rajan fails

to disclose or even remotely suggest at least several features of claim 33, including the feature of object descriptions generated by performing object extraction processing. Accordingly, because Rajan fails to disclose or suggest at least these claimed features, this reference fails to anticipate independent claim 33. Applicants respectfully submit that this claim is in condition for allowance. Additionally, because all depending claims 34-43 contain the foregoing limitations through dependency, Applicants respectfully submit that these claims are also in condition for allowance.

**CONCLUSION**

In view of the foregoing remarks, favorable consideration and allowance of claims 1-43 are respectfully solicited. In the event that the application is not deemed in condition for allowance, the examiner is invited to contact the undersigned in an effort to advance the prosecution of this application.

Respectfully submitted,



---

Paul A. Ragusa  
PTO Reg. No. 38,587

Robert L. Maier  
PTO Reg. No. 54,291

Attorneys for Applicants  
(212) 408-2538

BAKER BOTTS L.L.P.  
30 Rockefeller Plaza  
New York, NY 10112